

Curriculum Vitae et Studiorum

MARIA FEDERICO

Ottobre 2012

Informazioni Personali

- **Nome e Cognome:** Maria Federico
- **Luogo e Data di nascita:** Reggio Calabria, 12 Gennaio 1982
- **Indirizzo ufficio:** Dipartimento di Comunicazione ed Economia, Università di Modena e Reggio Emilia, Viale Allegrì 15, Reggio Emilia
- **Telefono:** 0522/523156
- **E-mail:** maria.federico@unimore.it, maria.federico@gmail.com
- **Home page:** <http://algo.ing.unimo.it/people/maria/index.html>

Posizione attuale

- **Luglio 2012 – Gennaio 2013:** Incarico di collaborazione coordinata e continuativa inerente al progetto “Realizzazione di uno sportello relativo all’utilizzo di strumenti dispensativi a favore degli studenti disabili e con Disturbo Specifico dell’Apprendimento (DSA) dell’Università degli Studi di Modena e Reggio Emilia” presso il Servizio Accoglienza Studenti Disabili dell’Università di Modena e Reggio Emilia
- Docente a contratto del corso *Web usability*, Corso di Laurea Magistrale in Pubblicità, Editoria e Creatività d’Impresa, Curriculum: Web communication, Dipartimento di Comunicazione ed Economia, Università di Modena e Reggio Emilia, per l’anno accademico 2012-2013 [2° semestre]
- Docente a contratto del corso *Laboratorio di grafica e videocomunicazione*, Corso di Laurea in Scienze della Comunicazione, Dipartimento di Comunicazione ed Economia, Università di Modena e Reggio Emilia, per l’anno accademico 2012-2013 [2° semestre]

Titoli di Studio

- **07 Marzo 2011:** Dottorato di Ricerca conseguito presso la Scuola di Dottorato in Modellistica, Simulazione Computazionale e Caratterizzazione Multiscala per le Scienze dei Materiali e della Vita, Università di Modena e Reggio Emilia.
Titolo tesi: Discovery of motifs and long multiple repeats in biological sequences
- **13 Ottobre 2006:** Laurea Specialistica in Tecnologie Informatiche, conseguita presso la Facoltà di Scienze Matematiche, Fisiche e Naturali, Università di Pisa.
Titolo tesi: Nozioni di massimalità per motivi in sequenze biologiche
Votazione: 110/110 e lode

- **05 Marzo 2004:** Laurea in Informatica, conseguita presso la Facoltà di Scienze Matematiche, Fisiche e Naturali, Università di Pisa.
Titolo tesi: Archivio delle pubblicazioni del Dipartimento di informatica
Votazione: 109/110
- **Luglio 2000:** Maturità scientifica conseguita presso il Liceo Scientifico XXV Aprile di Pontedera (PI)
Votazione: 85/100

Esperienze di Ricerca presso Istituti all'estero

Agosto 2009 – Aprile 2010: *Visiting Scholar* presso il BAOBAB team, guidato da Marie-France Sagot, dell'INRIA Rhône-Alpes e Laboratoire de Biométrie et Biologie Évolutive, Université Claude Bernard – Lyon 1 (France)

Attività di ricerca: Ricerca di ripetizioni multiple in sequenze biologiche usando tecniche di filtraggio

Competenze

Lingue straniere

- Ottima conoscenza dell'inglese scritto e parlato
- Buona conoscenza del francese scritto e parlato

Linguaggi di programmazione

- Ottima conoscenza di C, C++, Java, Perl
- Ottima conoscenza di Java applet, Javascript, JSP, PHP, (X)HTML, XML, CSS
- Ottima conoscenza di bash shell scripting

Attività Didattica

Corsi

[A.A. 2011-2012]

- *Web usability*, Corso di Laurea Magistrale in Pubblicità, Editoria e Creatività d'Impresa, Curriculum: Web communication, Facoltà di Scienze della Comunicazione e dell'Economia, Università di Modena e Reggio Emilia
- *Laboratorio di grafica e videocomunicazione*, Corso di Laurea in Scienze della Comunicazione, Facoltà di Scienze della Comunicazione e dell'Economia, Università di Modena e Reggio Emilia (in presenza e in modalità formazione a distanza (FAD))

[A.A. 2010-2011]

- *Web usability*, Corso di Laurea Magistrale in Pubblicità, Editoria e Creatività d'Impresa, Curriculum: Web communication, Facoltà di Scienze della Comunicazione e dell'Economia, Università di Modena e Reggio Emilia

Esercitazioni e Seminari Didattici

[A.A. 2011-2012]

- *Tecnologia della comunicazione*, Corso di Laurea in Scienze della Comunicazione, Facoltà di Scienze della Comunicazione e dell'Economia, Università di Modena e Reggio Emilia (periodo: 8/10/2011-16/12/2011, docente del corso: Dr Manuela Montangero)

[A.A. 2009-2010]

- *Algoritmi e Strutture Dati*, Corso di Laurea in Informatica, Facoltà di Scienze Matematiche Fisiche e Naturali, Università di Modena e Reggio Emilia (periodo: 30 ore, docente del corso: Prof. Mauro Leoncini)

Attività Scientifica e Professionale

Attività di ricerca presso Istituti in Italia e all'estero

- **Maggio 2011 – Marzo 2012:** Incarico di collaborazione coordinata e continuativa inerente al progetto “Analisi e progettazione di meccanismi per migliorare l'accessibilità dei contenuti multimediali in ambito didattico” presso il Servizio Accoglienza Studenti Disabili dell'Università di Modena e Reggio Emilia
- **Agosto 2009 – Aprile 2010:** *Visiting* presso il BAOBAB team, guidato da Marie-France Sagot, dell'INRIA Rhône-Alpes e Laboratoire de Biométrie et Biologie Évolutive, Université Claude Bernard – Lyon 1 (Francia)
- **Maggio 2007 – Dicembre 2007:** Assegno di Ricerca con tema “Modelli e Metodi per il Trattamento delle Lingue Naturali, e Prototipi Applicativi Mono e Multilingua”, presso l'Istituto di Linguistica Computazionale “Antonio Zampolli”, Consiglio Nazionale delle Ricerche (CNR) di Pisa
- **Febbraio 2007 – Marzo 2007:** Incarico di collaborazione professionale occasionale nell'ambito del progetto “SmartText: Progettazione e Sviluppo di un software per il chunking stocastico della lingua italiana basato sul principio della massimizzazione dell'entropia”, presso l'Istituto di Linguistica Computazionale “Antonio Zampolli”, Consiglio Nazionale delle Ricerche (CNR) di Pisa

Partecipazione a Comitati di Programma di Conferenze Internazionali

- International Conference on Information Technology in Bio- and Medical Informatics (ITBAM'12) in conjunction with the 23th DEXA – DEXA'12, 3-7 Settembre, 2012, Vienna (Austria). Membro dell'IPC.
- International Conference on Information Technology in Bio- and Medical Informatics (ITBAM'11) in conjunction with the 22nd DEXA – DEXA'11, 29 Agosto-2 Settembre, 2011, Tolosa (Francia). Membro dell'IPC.
- International Conference on Information Technology in Bio- and Medical Informatics (ITBAM'10) in conjunction with the 21st DEXA – DEXA'10, 30 Agosto-3 Settembre, 2010, Bilbao (Spagna). Membro dell'IPC.

Partecipazione a Progetti di Ricerca

- **2008-2009 Progetto FIRB:** Internazionalizzazione Italia-Israele on Pattern Discovery in Discrete Structures with Applications to Bioinformatics. Coordinatore Nazionale: Prof. Raffaele Giancarlo (Università di Palermo)
- **2008 PRIN:** Problemi inversi in medicina e astronomia. Coordinatore Scientifico: Prof. Mario Bertero (Università degli Studi di Genova)

Attività di Revisore per Conferenze

- International Conference on Information Technology in Bio- and Medical Informatics (ITBAM'12) in conjunction with the 23th DEXA – DEXA'12, 3-7 Settembre, 2012, Vienna (Austria).
- International Conference on Information Technology in Bio- and Medical Informatics (ITBAM'11) in conjunction with the 22nd DEXA – DEXA'11, 29 Agosto – 2 Settembre, 2011, Tolosa (Francia).
- International Conference on Information Technology in Bio- and Medical Informatics (ITBAM'10) in conjunction with the 21st DEXA – DEXA'10, 30 Agosto – 3 Settembre, 2010, Bilbao (Spagna).

Relazioni tenute a Congressi e Convegni Nazionali e Internazionali

- *Enhancing Learning Accessibility through Fully Automatic Captioning*. 9th International Cross-Disciplinary Conference on Web Accessibility – W4A 2012, Lione (Francia), 16-17 Aprile, 2012.
- *Removing artifacts of approximated motifs*. 2nd International Conference on Information Technology in Bio- and Medical Informatics - ITBAM 2011, Tolosa (Francia), 29 Agosto – 2 Settembre, 2011.
- *An optimized filter for finding multiple repeats in DNA sequences*. 8th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA 2010), Hammamet (Tunisia), 16-19 Maggio, 2010.
- *An Efficient Algorithm for Planted Structured Motif Extraction*. Workshop on Breaking Frontiers of Computational Biology (COMPBIO'09), Ischia (Italia), 18-20 Maggio, 2009.
- *Suffix Tree Characterization of Maximal Motifs in Biological Sequences*. FIRB'09 Annual Project Meeting for Pattern Discovery in Discrete Structures with Applications to Bioinformatics, Università di Palermo, Dipartimento di Matematica e Applicazioni (Italia), 23-24 Febbraio, 2009
- *Suffix Tree Characterization of Maximal Motifs in Biological Sequences*. Workshop on Algorithms in Molecular Biology (ALBIO'08), Vienna (Austria), 7-9 Luglio, 2008
- *Maximum Entropy for Italian PoS Tagging*. EVALITA 2007, Roma (Italia), 10 Settembre, 2007

Elenco pubblicazioni

Articoli su Riviste Nazionali e Internazionali

- [1] M. Federico, M. Leoncini, M. Montangero, P. Valente. “Direct vs 2-Stage Approaches to Structured Motif Finding”, *Algorithms for Molecular Biology*, 2012 Aug 21;7(1):20. [Epub ahead of print]. ISSN: 1748-7188
<http://www.almob.org/content/pdf/1748-7188-7-20.pdf>
- [2] N. Dusi, M. Federico, M. Furini. “MovieRemix: Having Fun Playing with Videos”, *International Journal of Computer Games Technology*, vol. 2011, Article ID 857371, 13 pages, 2011. doi:10.1155/2011/857371
<http://downloads.hindawi.com/journals/ijcgt/2011/857371.pdf>
- [3] M. Federico, N. Pisanti. “Suffix Tree Characterization of Maximal Motifs in Biological Sequences”, *Theoretical Computer Science*, Vol. 410, Issue 43, pages 4391-4401, 2009. ISSN:0304-3975
<http://www.sciencedirect.com/science/article/pii/S0304397509004812>
- [4] F. Dell'Orletta, M. Federico, A. Lenci, S. Montemagni, V. Pirrelli. 2007. “Maximum Entropy for Italian PoS Tagging”. *Intelligenza artificiale*, Vol. IV, Numero 2, pag. 10-11, Giugno 2007. ISSN: 1724-8035
<http://www.evalita.it/sites/evalita.fbk.eu/files/proceedings2007/03-IA-IV-2-ILCcnrUniPi-pos.pdf>

Articoli in Atti di Congresso Internazionali

- [5] M. Federico, M. Furini. “Enhancing Learning Accessibility through Fully Automatic Captioning”, *Proceedings of the 9th International Cross-Disciplinary Conference on Web Accessibility*, Lione (Francia), 16-17 Aprile, 2012. ISBN: 978-1-4503-1019-2
<http://dl.acm.org/citation.cfm?id=2207053>
- [6] M. Federico, P. Peterlongo, N. Pisanti, M.-F. Sagot. “Finding Long and Multiple Repeats with Edit Distance”, *Proceedings of the Prague Stringology Conference 2011*, pag. 83-97, Praga (Repubblica Ceca), 29-31 Agosto, 2011. ISBN: 978-80-01-04870-2
<http://www.stringology.org/cgi-bin/getfile.cgi?t=pdf&c=-&y=2011&n=08>
- [7] M. Federico, N. Pisanti. “Removing artifacts of approximated motifs”, *Proceedings of the 2nd International Conference on Information Technology in Bio- and Medical Informatics - ITBAM 2011*, Lecture Notes in Computer Science (LNCS) by Springer Verlag, Tolosa (Francia), 29 Agosto-2 Settembre, 2011. ISBN: 978-3-642-23207-7
<http://dl.acm.org/citation.cfm?id=2035504>
- [8] L. Palmieri, M. Federico, M. Leoncini, M. Montangero, P. Valente. “A High Performing Tool for Residue Solvent Accessibility Prediction Using Sequence Homology”. *Proceedings of the 2nd International Conference on Information Technology in Bio- and Medical Informatics - ITBAM 2011*, Lecture Notes in Computer Science (LNCS) by

Springer Verlag, Tolosa (Francia), 29 Agosto-2 Settembre, 2011. ISBN: 978-3-642-23207-7

<http://dl.acm.org/citation.cfm?id=2035503>

- [9] M. Federico, P. Peterlongo, N. Pisanti. “An optimized filter for finding multiple repeats in DNA sequences”, *Proceedings of the 8th ACS/IEEE International Conference on Computer Systems and Applications (AICCSA 2010)*, Hammamet (Tunisia), 16-19 Maggio, 2010. ISBN: 978-1-4244-7715-9

http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5587026&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D5587026

- [10] M. Federico, P. Valente, M. Leoncini, M. Montangero, R. Cavicchioli. “An Efficient Algorithm for Planted Structured Motif Extraction”, *Proceedings of ACM International Conference on Computing Frontiers 2009 (COMPBIO'09)*, Ischia (Italia), 8-20 Maggio, 2009. ISBN:978-1-60558-556-7

<http://dl.acm.org/citation.cfm?id=1531782>

- [11] M. Federico, N. Pisanti. 2008. “Suffix Tree Characterization of Maximal Motifs in Biological Sequences”, *Proceedings of Bioinformatics Research and Development (BIRD)*, Springer Communications in Computer and Information Science 13, pages 456--465, Vienna (Austria), 7-9 Luglio, 2008. ISBN: 978-3-540-70598-7

<http://www.sciencedirect.com/science/article/pii/S0304397509004812>

- [12] A. Cappelli, E. Giovannetti, M. Federico. 2005. “The Role of Communication in Human-Robot Interaction”, *Proceedings of the Second Robocare Workshop*, pages 19-24, Roma (Italia), 5 Maggio, 2005. ISBN 88-85059-19-8

<http://robocare.istc.cnr.it/RC-Ws-2-proceedings.pdf>

Rapporti Tecnici

- [13] F. Dell'Orletta, M. Federico, E. Giovannetti, A. Lenci, S. Marchi, A. Trabucco, V. Pirrelli. “Language Recognition Tool, Specifiche di Implementazione”, Rapporto Tecnico ILC TW 2007 01- Progetto Smarttext, ILC-CNR, Pisa, Italia, 2007.

- [14] F. Dell'Orletta, M. Federico, E. Giovannetti, A. Lenci, S. Marchi, A. Trabucco, V. Pirrelli. “Segmentazione di un Testo Italiano in Token”, Rapporto Tecnico ILC TW 2007 02- Progetto Smarttext, ILC-CNR, Pisa, Italia, 2007.

- [15] F. Dell'Orletta, M. Federico, E. Giovannetti, A. Lenci, S. Marchi, A. Trabucco, V. Pirrelli. “Analisi Morfosintattica per l'Italiano”, Rapporto Tecnico ILC TW 2007 03- Progetto Smarttext, ILC-CNR, Pisa, Italia, 2007.

- [16] F. Dell'Orletta, M. Federico, E. Giovannetti, A. Lenci, S. Marchi, A. Trabucco, V. Pirrelli. “Part-of-Speech Tagging per l'Italiano”, Rapporto Tecnico ILC TW 2007 04- Progetto Smarttext, ILC-CNR, Pisa, Italia, 2007.

- [17] F. Dell'Orletta, M. Federico, E. Giovannetti, A. Lenci, S. Marchi, A. Trabucco, V. Pirrelli. “Specifiche di Chunking per l'Italiano”, Rapporto Tecnico ILC TW 2007 05- Progetto Smarttext, ILC-CNR, Pisa, Italia, 2007.

- [18] F. Dell'Orletta, M. Federico, E. Giovannetti, A. Lenci, S. Marchi, A. Trabucco, V. Pirrelli. "Specifiche di Named Entity Recognition per l'Italiano", Rapporto Tecnico ILC TW 2007 06- Progetto Smarttext, ILC-CNR, Pisa, Italia, 2007.

Articoli sottomessi

- M. Federico, M. Furini. "An Automatic Caption Alignment Mechanism for off-the-shelf Speech Recognition Technologies". Sottomesso a rivista internazionale.
- M. Federico, P. Peterlongo, N. Pisanti, M.-F. Sagot. "RIME: Repeat Identification". Sottomesso a rivista internazionale.

Abstract tesi di dottorato

In this thesis, we address the problem of the discovery of both short and long repeated patterns in biological sequences. For what concerns short repetitions, we focus on exact methods for the extraction of structured motifs approximated with the Hamming distance. We perform a comparison of different approaches to this problem both on synthetic and biological data in order to deduce a set of conditions on parameter sets which should allow biologists to choose the more suitable tool for their needs.

The output size of algorithms to find motifs might be very large (namely exponential in the approximation degree) and redundant, making the output often unreadable, next to slow down the inference itself. We aim at removing the redundancy in the output mainly due to nested motifs, and hence, in this thesis, we extend some notions of maximality already defined for exact motifs to the case of approximate motifs with Hamming distance, and we give a characterization of maximal motifs on the suffix tree. Given that this data structure is used by a whole class of motif extraction tools, we show how these tools can be modified to include the maximality requirement without changing the asymptotical complexity.

For what concerns long repetitions, we address the problem of finding long repeats having two or more occurrences in a set of sequences. This problem is computationally challenging when a certain non negligible number of insertions, deletions and substitutions are admitted among repeat occurrences. We present an ab initio method, called MRF, which finds long repeats in one or more sequences taking advantage of a filtering technique which discards as many as possible fragments of the sequences that are guaranteed not to contain any searched repeat. The information gathered by the filter is then used in a successive inference phase performed on filtered data. To the best of our knowledge, MRF is the first tool that can deal with repeats occurring possibly several times, that have length of hundreds or thousands of bases, and whose occurrences may differ in even more than 10% of their positions in terms of substitutions and indels.

La sottoscritta è a conoscenza che, ai sensi dell'art. 26 della legge 15/68, le dichiarazioni mendaci, la falsità negli atti e l'uso di atti falsi sono puniti ai sensi del codice penale e delle leggi speciali. Inoltre, la sottoscritta autorizza al trattamento dei dati personali, secondo quanto previsto dal D.Lgs. n. 196/2003 del 30 giugno 2003 e successive modifiche ed integrazioni, per le finalità di cui al presente avviso di candidatura.

Maria Federico